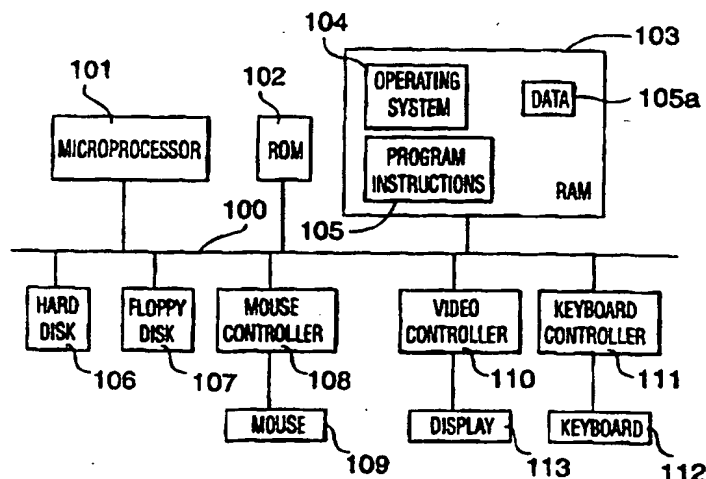




## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>G06F 17/30</b>	<b>A1</b>	(11) International Publication Number: <b>WO 98/09229</b> (43) International Publication Date: <b>5 March 1998 (05.03.98)</b>
<p>(21) International Application Number: <b>PCT/CA97/00611</b></p> <p>(22) International Filing Date: <b>29 August 1997 (29.08.97)</b></p> <p>(30) Priority Data: <b>2,184,518</b>      <b>30 August 1996 (30.08.96)</b>      <b>CA</b></p> <p>(71) Applicant (for all designated States except US): <b>TELEVITESSE SYSTEMS INC. [CA/CA]; 603 March Road, Kanata, Ontario K2K 2M5 (CA).</b></p> <p>(72) Inventors; and (75) Inventors/Applicants (for US only): <b>REED, Jim [CA/CA]; 26 Sheperds Glen Avenue, Kanata, Ontario K2M 2M9 (CA). STREATCH, Paul [CA/CA]; P.O. Box 1196, Richmond, Ontario K0A 2Z0 (CA).</b></p> <p>(74) Agent: <b>MITCHELL, Richard, J.; Marks &amp; Clerk, P.O. Box 957, Station B, Ottawa, Ontario K1P 5S7 (CA).</b></p>		<p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW.</b> ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).</p> <p><b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>

(54) Title: **REAL TIME STRUCTURED SUMMARY SEARCH ENGINE**

## (57) Abstract

A method of organizing electronic documents for storage and subsequent retrieval, involves storing a summary structure describing the structure of summary records associated with each document. Each structured summary record has at least one field representative of a characteristic of the document. A predetermined number of field values identify the value of the characteristic associated with the field. Predetermined keyword criteria associated with the field values are stored. Each document is analyzed to build a text index listing the occurrence of unique significant words in the document. The text index is compared with the keyword criteria to determine the appropriate field value for the document. For example, one characteristic field might be related to topic, which could have the field values of "financial" or "sports". The preponderance of certain keyword criteria, such as "money" or "shares" would identify the document with the financial topic.

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

- 1 -

## REAL TIME STRUCTURED SUMMARY SEARCH ENGINE

This invention relates to a method of processing data, and more particularly to a method of processing stored electronic documents to facilitate subsequent retrieval.

It is known to search text-based documents electronically using keywords linked through Boolean logic. This technique has been used for many years to search patent literature, for example, and more recently documents on the Internet. The problem with such conventional searches is that if the search criteria are made broad, the search engine will often produce thousands of "hits", many of which are of no interest to the searcher. If the criteria are made too narrow, there is a risk that relevant documents will be missed.

There is a real need to provide a search engine that will filter out unwanted results while retaining results of interest to the user. An object of the invention is to provide such a system.

According to the present invention there is provided a method of processing electronic documents for subsequent retrieval, comprising the steps of storing in memory a summary structure describing the structure of summary records associated with each document, each structured summary record having at least one field representative of a characteristic of the document and having a predetermined number of field values identifying the value of the characteristic associated therewith; storing in memory predetermined keyword criteria associated with said field values; analyzing each document to build a text index listing the occurrence of unique significant words in the document; and comparing said text index with said keyword criteria to determine the appropriate field value for the document.

Examples of fields with limited field values are category and location. The category field might have as possible field values: Finance, Sports, Politics. The location field might have as possible values: Africa, Canada, Europe.

The individual field values are in turn associated with certain keyword criteria. For example, the criteria for the financial field value might be: shares, public, bankrupt, market, profit, investor, stock, IPO, quarter, "fund manager". The criteria for the sports field value might be: football, ball, basketball, hockey, bat, score, soccer, run, baseball, "Wayne Gretsky", "Chicago Bulls", "Michael Jordan".

It will be appreciated that the keyword criteria are chosen in view of the likelihood that any document containing those keywords will be associated with the particular category.

- 2 -

In a preferred embodiment, the structured summary also includes fields having unlimited values. Examples of such fields are a keyword field and an excerpt field. The keyword field may list the words having the highest count in the text index. The excerpt field may list the sentences containing the highest occurrence of keywords.

5           The structured summary can be established according to a standard profile that is the same for all users, or in one embodiment the profile can change in accordance with a particular user's need. In this case, a user profile is stored in a profile database.

10           The structured summaries normally include pointers to the memory locations of the associated documents so that during a subsequent search, a user view relevant summaries and quickly locate the associated document as required.

15           The invention also extends to a system for processing electronic documents for subsequent retrieval, comprising a memory storing a summary structure describing the structure of summary records associated with each document, each structured summary record having at least one field representative of a characteristic of the document and having a predetermined number of field values identifying the value of the characteristic associated therewith; a memory storing predetermined keyword criteria associated with said field values; means for analyzing each document to build a text index listing the occurrence of unique significant words in the document; and means for comparing said text index with said keyword criteria to determine the appropriate field value for the document.

20           The invention still further provides a method of retrieving electronic documents which are associated with a structured summary record containing a pointer to the document and having at least one field representative of a characteristic of the document and having a predetermined number of field values identifying the value of the characteristic associated therewith, comprising searching through the summary records for records having specific field values, and identifying the documents associated with the records matching the search criteria.

25           The invention will now be described in more detail, by way of example, only with reference to the accompanying drawings, in which:-

30           Figure 1 is a block diagram of a computer system suitable for implementing the present invention;

            Figure 2 is data flow diagram for a method in accordance with the invention; and

- 3 -

Figure 3 is a flow chart illustrating the operation of a part of the method in accordance with the invention.

- The following table is an example of a structured summary record associated with a particular document, in this case an article on the Internet search engine, Yahoo. The record has two limited fields, category and location, having, for example, the field values, for example, finance, sports, and politics for category, and Africa, Canada, and Europe for location, and two fields, keywords and excerpts, having unlimited field values.

STRUCTURED SUMMARY

Field Type	Field Value
Category	Financial
Location	Canada
Keywords	Yahoo, Internet. Search, Software
Excerpts	Shares in the maker of Internet search software are tumbling. Yahoo stock (YHOO/NASDAQ) is down 38% from April's first-day trading high of US\$33 as investors pull out on fears of increasing competition and lack of proprietary technology.

- In this example, the value for category is *financial* and the value for location is *Canada*. The unlimited fields contain keywords and key sentences, i.e. sentences containing the highest occurrence of keywords.

- The structured summary records for a series of documents are stored in a database, for example, on a computer hard disk as a series of such records, each having a pointer to the location in memory of the associated document that it summarizes. When a user wishes to perform a search, he or she can search through the structured summaries, for example, for the keyword *Yahoo*, looking only for those records that have the field value *financial* for category.

- Each limited field value contains a pointer to another entry in a database of summary candidate databases. Each record in this database identifies the keyword criteria associated with each field value of the structured summary record. Each candidate has a name corresponding to a field value of the structured summary record. The table below illustrates a summary candidate database. The first record has a candidate name *financial*, which is one of the values for the field name *category* in the structured summary. The

- 4 -

candidate *financial* lists the keywords that identify a documents as belonging to the category *financial*.

## SUMMARY CANDIDATE DATABASE

Field Name	Candidate name	Keyword Criteria
Category	Financial	shares, public, bankrupt, market, profit, investor, stock, IPO, quarter, "fund manager
Category	Sports	football, ball, basketball, hockey, bat, score, soccer, run, baseball, "Wayne Gretsky", "Chicago Bulls", "Michael Jordan".
Location	Canada	Canada, Toronto, Ottawa, Vancouver, Halifax etc.
Location	Asia	Asia, Far East, Japan, Tokyo, Korea, etc.
Location	Europe	Europe, London, Paris, Germany etc.

5 A plurality of summary structures can be stored in the summary structure database in accordance with the user profile and each such structure is given a unique name to identify the particular user or class of users.

10 The invention is implemented on a general purpose computer, such as an IBM-compatible Pentium-based personal computer, although more powerful computers can be employed to increase storage capacity and decrease search time. The summary candidate database and the structured summaries can be stored on a hard disk.

15 In order to implement the invention, a computer as shown in Figure 1 is loaded with instructions to implement the described method. The computer typically comprises a main bus 100 connected to a microprocessor 101, a Read-Only-Memory 102 and a Random Access Memory 103 containing the operating system 104, program instructions 105 implementing the present invention, and data 105a.

20 Typically connected to the bus 100 are a hard disk 106, which stores the database information, a floppy disk 107, a mouse controller 108 connected to mouse 109, a video controller 110 connected to display 113, and a keyboard controller 111 connected to a keyboard 112.

- 5 -

As shown in Figure 2, the computer first reads the structured summary database 1 stored on hard disk 106 to extract the summary structure 2. This can be made user dependent, or alternatively can be the same for all users. The summary structure record contains the field structure of the summary records to be created. The system then extracts  
5 the next electronic document from a document database 3 and builds 4 a text index 5, which is temporarily stored in memory. This consists of an index of all significant words in the document, i.e. excluding "noise words", such as "or", "and", "the" etc. and ranks them according to word count.

10 The computer then generates a at step 6, structured summary 7 which is stored in memory 103.

A detailed flow chart illustrating the generation of each summary record is shown in Figure 3. At the start 10, the system creates a new summary record 11 associated with a new document extracted from the document database. The new record has a field structure defined in the field structure database and includes a pointer to the memory  
15 location of the associated document. During operation of the loop, the system keeps track in memory of the name of a "current candidate" and its word count (to be described). At block 11, the system is also initialized to set the *current candidate* and corresponding *word count* to none.

At step 12, the system sets the summary record field name to the next unique field  
20 name in the summary structure database starting from the first, and at 13 retrieves from the summary candidate database the next summary candidate (selected candidate) also starting from the first having a field name matching the summary record field name that has just been set. For example, the first summary record field name might be "category". The first summary candidate with a field name category might be "financial" having the  
25 criteria keywords noted above.

Next, the number of occurrences of each word on the criteria word list in the current document for the selected candidate (financial) is determined at 14 and these occurrences are totaled to give the *word count* for the selected candidate. Decision unit 15  
30 determines whether the total word count for the selected candidate is greater than the word count for the current candidate. If the answer is yes, the current candidate is set to the selected candidate. Clearly, on the first pass, the current candidate will be set to the selected candidate unless none of the criteria keywords appear in the document.

Decision unit 17 determines whether there are any more candidate records in the candidate database, and if so the loop is repeated for the next candidate. Decision unit 15

- 6 -

determines whether the candidate word count is greater than the word current of the current candidate, and if so unit 16 sets the new selected candidate to the current candidate. Otherwise the loop is repeated until there are no more candidates, whereupon the summary field value of the structured summary is set to the name of the current candidate at unit 18.

The larger loop is repeated 19 until there are no more field names. The net result is that the structured summary contains a series of field names which have values corresponding to the names of the summary candidates whose word count is the highest for the corresponding field name, unless of course none of the keywords for any of the values of a particular field name appear in the target document, in which case the field value will remain blank.

In a preferred implementation, when the summary structure database is first read, an index is built that maps the words contained in the criteria word lists to summary candidates. With this arrangement, it is easy to determine a sub-set of summary candidates that are applicable to the current document. By counting only the words in the summary candidates that are applicable, summaries using a large summary database (>100,000 criteria words) can be quickly generated. The use of a large summary database is the key to generating accurate summaries.

A similar loop determines the keywords having the highest count, for example, the first four, and enters these into the keyword field. Another loop determines the sentences, for example, containing the keywords having the highest count, for example, the first four sentences with the highest occurrence of keywords.

The described real-time structured summary system provides a technology can be used as the basis for developing, a number of sophisticated search features that will help the user filter out unrelated results and focus on the results that are of interest.

The real power of having a structured summary is observed when a user summarizes a set of related documents, rather than just a single document (e.g., a set of clips accesTV™ Assistant, or a set of documents returned from an Internet Search Server. For example, a search for documents on Michael Jordon would return a hit from many documents of little interest to the user. If the results of the search are summarized, then the user can easily ignore stories that have, for example, the field *category* with a value other than *sports*.



- 7 -

Typically, a news story is re-broadcast many times throughout the day. Duplicate stories can be filtered out by comparing the summaries of recorded stories. If the summaries are the same, then there is a good chance that the documents are the same. This results in opening many fewer documents for comparison, which can be more efficient than the alternative.

It is also possible to use the system to look for similar documents. Predetermined criteria indicative of a degree of similarity can be set. For example, documents can be regarded as similar if there is a 90% match of keywords. In a search, the system can be asked to generate all summaries where there is a match of 90% or greater.

The system can be used with e-mail articles or news stories.

In another situation, consider the case of an Internet Search that has returned 3000 results, and a user has found a document that is of interest to them. The user can be presented with a short list of similar documents (hopefully much smaller than the 3000) using an application that looks for summaries (in the set of 3000 documents) that are similar to (have several fields in common) the summary of the document of interest.

An extension to determining similar documents, is an ignore feature. A user may be interested in monitoring stories on the Canadian Government, but not interested in continually receiving updates of Sheila Copps resigning. This feature can be implemented in the same manner as looking for similar documents, by looking for summaries that are similar to the summary of the document that is to be ignored.

Another feature allows a user to take a document that they may have received by e-mail, or downloaded from the Internet, and convert it to a search that can be used to monitor an accessTV Assistant source (e.g., Television channel), or that can be executed by an Internet Search Server. This feature can be implemented easily using summary technology. One possible implementation would be to monitor the summaries rather than the source, and look for similar summaries.

By adding a priority weight to summary items, it becomes very easy to prioritize results based on the user's individual interests. Results containing summary items with a higher weight will be given precedence over results with a lower weight.

- 8 -

By adding hierarchy information to summary items, a more sophisticated summary engine can be implemented. For example, a user might specify that a field type sub-category is dependent on a field type category, and that a particular sub-category named "basketball" is only applicable is the selected category is "sports". This way user can  
5 have a category hierarchy that results in a very accurate summary.

This embodiment could be applied to an automatic classification system for patent searching. Keywords most likely associated with particular classes and subclasses would need to be identified, and then the system would create structured summaries based on the highest occurrence of keywords. The system would enable documents to be retrieved  
10 electronically even though the entered keyword was not present in the document. For example, if the user was interested in a document relating to a microprocessor, but which did not actually use that term, the structured summary would contain the word *microprocessor* as a field value based on the fact that the document contained a high incidence of terms such as "program", "instruction", "processor" commonly associated  
15 with the term microprocessor.

The summary structures can also include a ranking field which keeps count of the number of relevant keywords, and this can be used to rank search results in order of importance.

- 9 -

## Claims:

1. A method of processing electronic documents for subsequent retrieval, comprising the steps of storing in memory a summary structure describing the structure of summary records associated with each document, each structured summary record having at least one field representative of a characteristic of the document and having a predetermined number of field values identifying the value of the characteristic associated therewith; storing in memory predetermined keyword criteria associated with said field values; analyzing each document to build a text index listing the occurrence of unique significant words in the document; and comparing said text index with said keyword criteria to determine the appropriate field value for the document.
2. A method as claimed in claim 1, wherein the appropriate field value is determined according to the keyword criteria having the highest count.
3. A method as claimed in claim 1, wherein said summary structure includes additional fields having unlimited values.
4. A method as claimed in claim 3, wherein one said additional field comprises a keyword field listing the words in said text index having the highest count.
5. A method as claimed in claim 3, wherein one said additional field comprises an excerpt field listing the sentences in said document containing the words in said text index with the highest count.
6. A method as claimed in claim 2, wherein a series of summary candidates corresponding to the field values of each characteristic field of the summary records are stored in memory, and said summary candidates are examined on an iterative basis to determine the summary candidate having the highest word count for the field.
7. A method as claimed in claim 6, further comprising building an index that maps the words contained in the criteria lists to the summary candidates so as to permit the subset of candidates applicable to a document to be rapidly determined.
8. A method as claimed in claim 1, wherein structured summaries are compared to determine whether two documents have the same content.
9. A method as claimed in claim 1, wherein structured summaries are compared to determine whether two documents have similar content based on a predetermined match of field values in the documents.
10. A method as claimed in claim 1, wherein said limited fields have at least one sub-field arranged in a hierarchical structure.

- 10 -

11. A method as claimed in claim 1, wherein said documents comprises news articles.
12. A method as claimed in claim 11, wherein said news articles are extracted from television broadcasts.
13. A method as claimed in claim 1, wherein said structured summary records also  
5 include a ranking field containing a keyword count to permit search hits to be ranked in order of importance.
14. A method as claimed in any one of claims 1 to 13, wherein each summary record contains a pointer to the location of the associated document.
15. A system for processing electronic documents for subsequent retrieval, comprising  
10 a memory storing a summary structure describing the structure of summary records associated with each document, each structured summary record having at least one field representative of a characteristic of the document and having a predetermined number of field values identifying the value of the characteristic associated therewith; a memory storing predetermined keyword criteria associated with said field values; means for  
15 analyzing each document to build a text index listing the occurrence of unique significant words in the document; and means for comparing said text index with said keyword criteria to determine the appropriate field value for the document.
16. A method of retrieving electronic documents which are associated with a structured summary record containing a pointer to the document and having at least one  
20 field representative of a characteristic of the document and having a predetermined number of field values identifying the value of the characteristic associated therewith, comprising searching through the summary records for records having specific field values, and identifying the documents associated with the records matching the search criteria.
- 25 17. A method as claimed in claim 16, further comprising searching for a specific keyword in said records matching the search criteria.

1/2

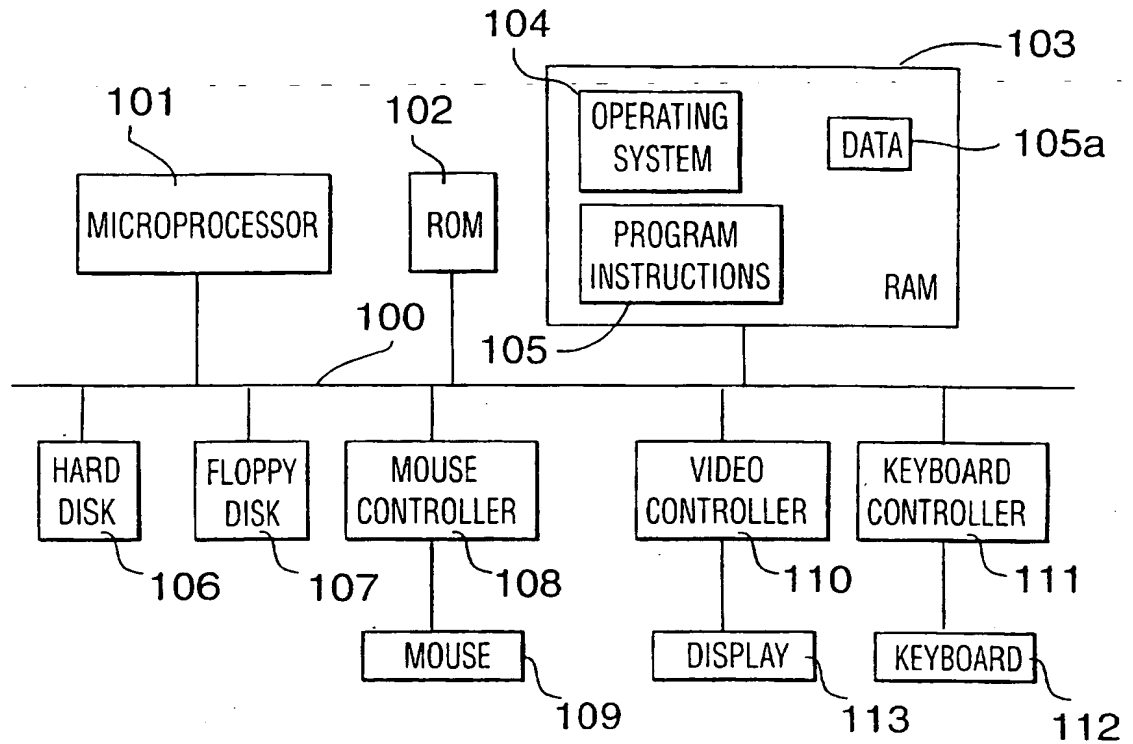


FIG. 1

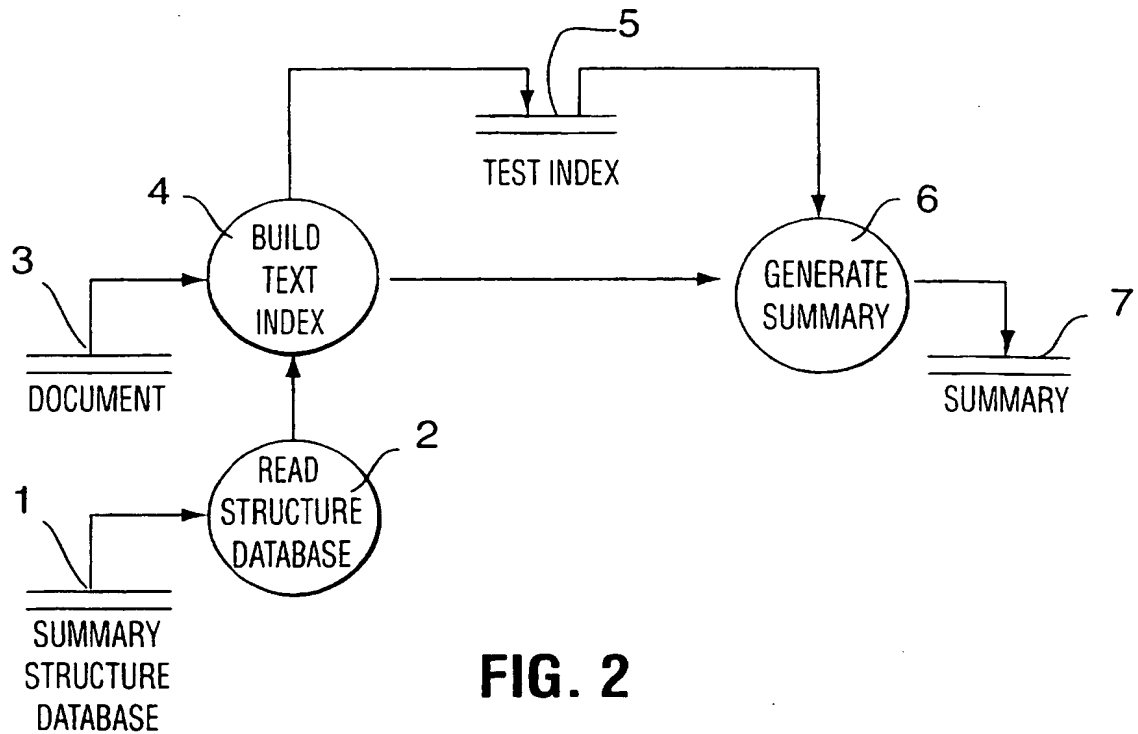
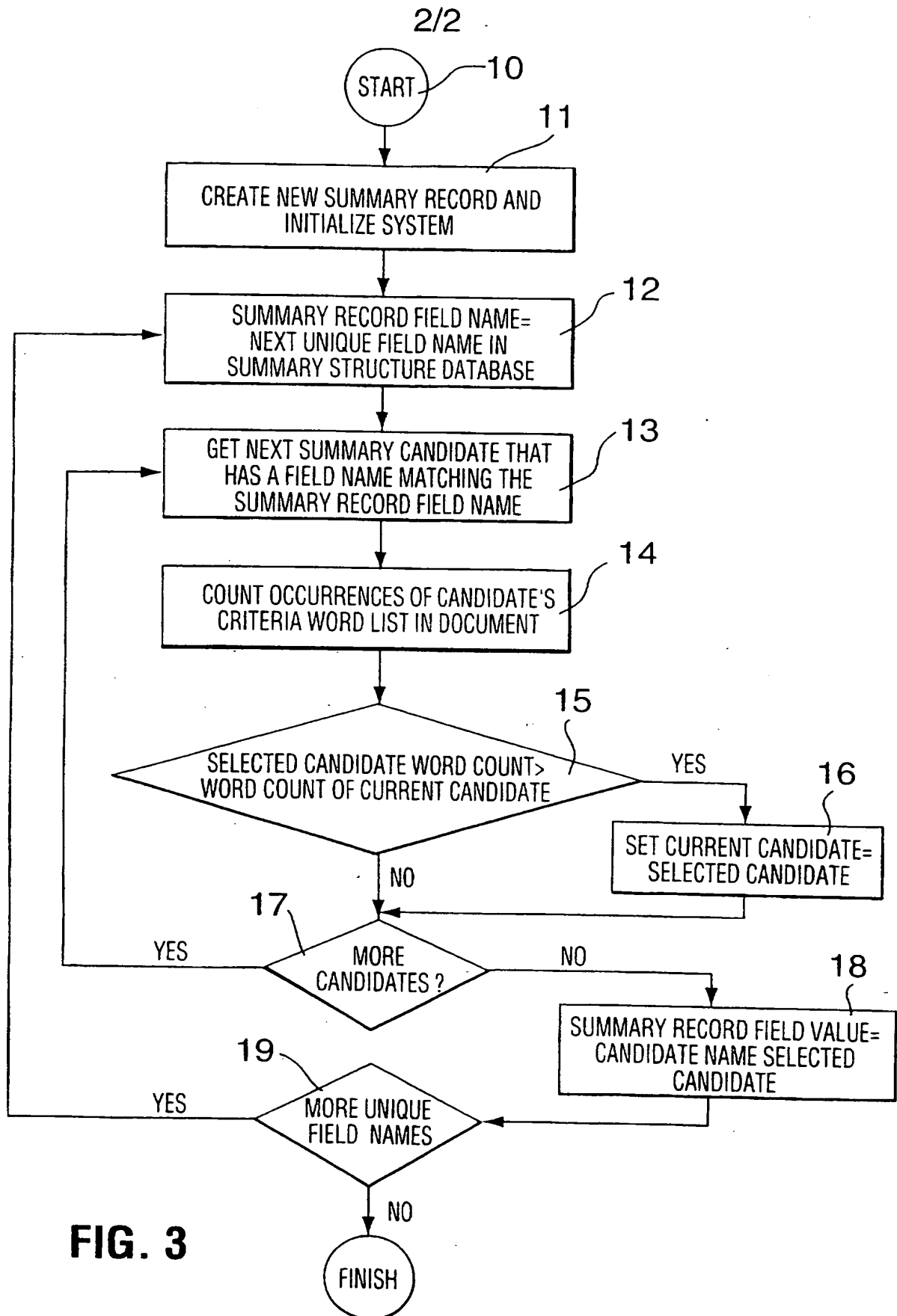


FIG. 2

SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)

# INTERNATIONAL SEARCH REPORT

Inter      national Application No  
PCT/CA 97/00611

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 6      G06F17/30

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols):  
IPC 6      G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
A	WO 96 23265 A (BRITISH TELECOMM ; DAVIES NICHOLAS JOHN (GB); WEEKS RICHARD (GB)) 1 August 1996 see page 4, line 18 - page 8, line 24 see page 13, line 17 - page 14, line 30 ---	1-17
A	JACOBS P S ET AL: "SCISOR: EXTRACTING INFORMATION FROM ON-LINE NEWS" COMMUNICATIONS OF THE ASSOCIATION FOR COMPUTING MACHINERY, vol. 33, no. 11, 1 November 1990, pages 88-97, XP000173090 see the whole document ---	1-17

-/--

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

16 December 1997

Date of mailing of the international search report

22/12/1997

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Fournier, C

# INTERNATIONAL SEARCH REPORT

Inter      nal Application No  
PCT/CA 97/00611

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication where appropriate, of the relevant passages	Relevant to claim No
A	<p>BEERUD SHETH ET AL: "EVOLVING AGENTS FOR PERSONALIZED INFORMATION FILTERING" PROCEEDINGS OF THE CONFERENCE ON ARTIFICIAL INTELLIGENCE FOR APPLICATIONS, ORLANDO, MAR. 1 - 5, 1993, no. CONF. 9, 1 March 1993, INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS, pages 345-352, XP000379626 see page 340, left-hand column, paragraph 3 - page 349, right-hand column, paragraph 5; figures 2-4</p> <p style="text-align: center;">---</p>	1-17
A	<p>PAICE C D: "CONSTRUCTING LITERATURE ABSTRACTS BY COMPUTER: TECHNIQUES AND PROSPECTS" INFORMATION PROCESSING &amp; MANAGEMENT (INCORPORATING INFORMATION TECHNOLOGY), vol. 26, no. 1, 1990, pages 171-186, XP000605124 see page 171, paragraph 1 - page 175, paragraph 3</p> <p style="text-align: center;">---</p>	1-17
A	<p>"METHOD FOR AUTOMATIC EXTRACTION OF RELEVANT SENTENCES FROM TEXTS" IBM TECHNICAL DISCLOSURE BULLETIN, vol. 33, no. 6A, November 1990, page 338/339 XP002015802 see the whole document</p> <p style="text-align: center;">-----</p>	

Form PCT/ISA/210 (continuation of second sheet) (July 1992)



# INTERNATIONAL SEARCH REPORT

Inter nal Application No

PCT/CA 97/00611

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9623265 A	01-08-96	AU 4454996 A	14-08-96
		CA 2210581 A	01-08-96
		EP 0807291 A	19-11-97
		FI 973080 A	22-07-97
		NO 973372 A	22-09-97
<hr/>			

Form PCT/ISA/210 (patent family annex) (July 1992)



1/2

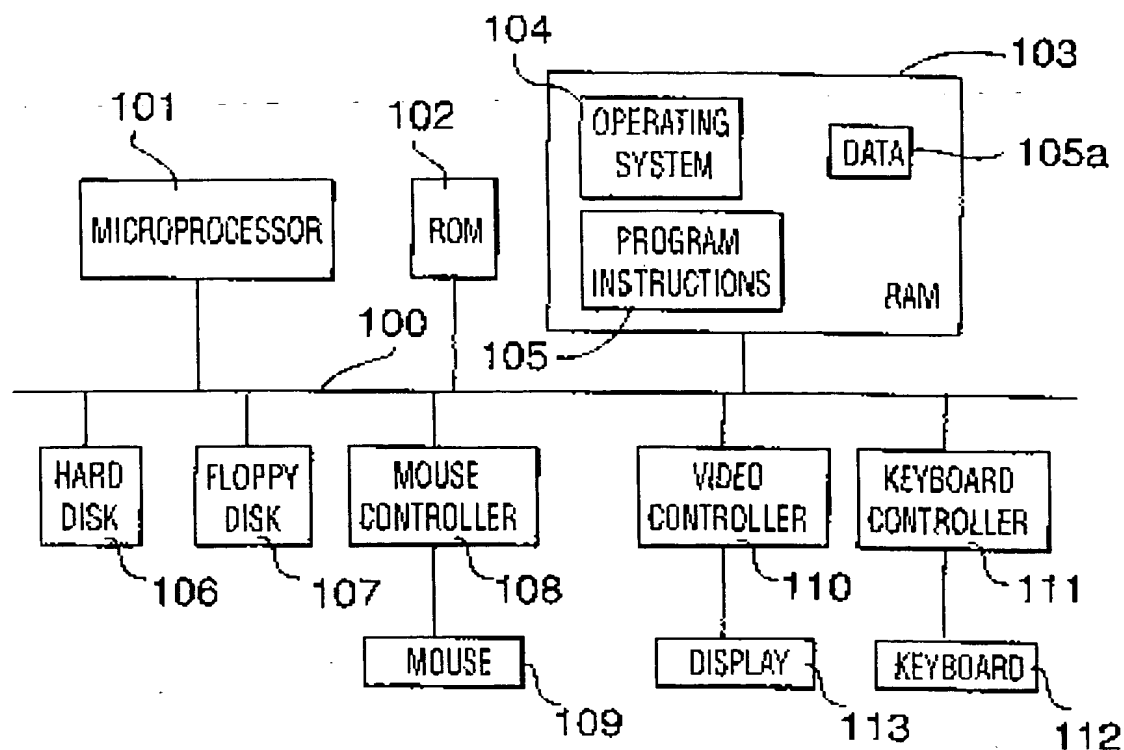


FIG. 1

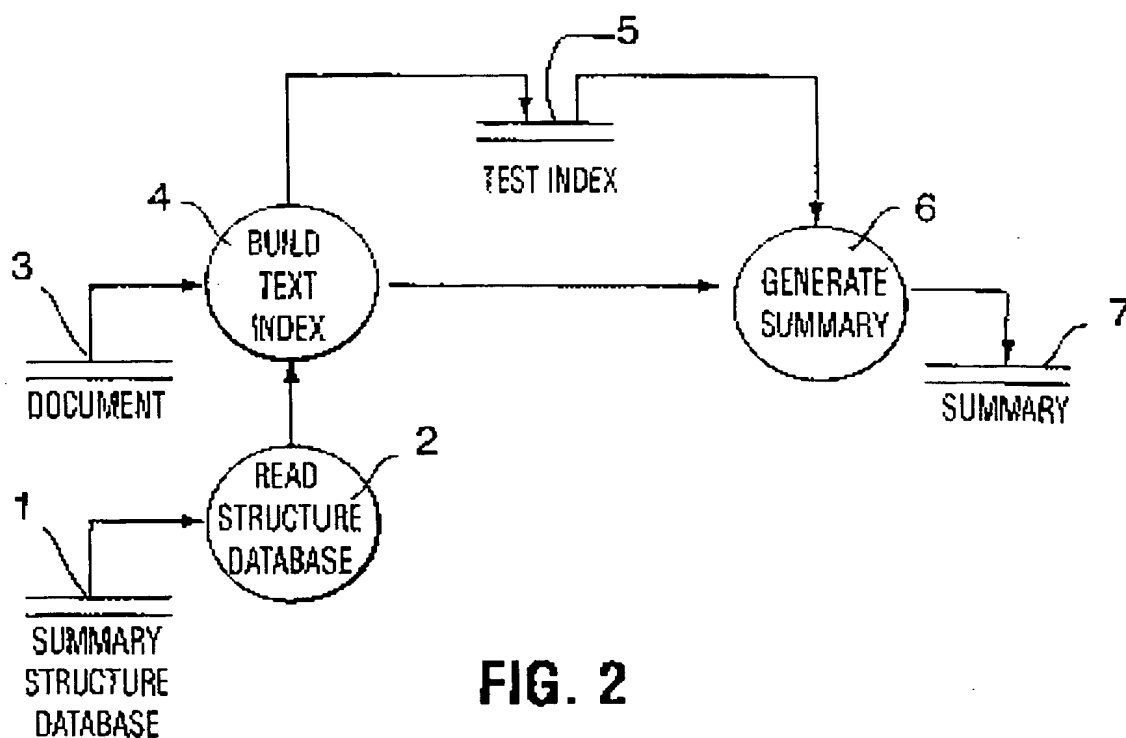
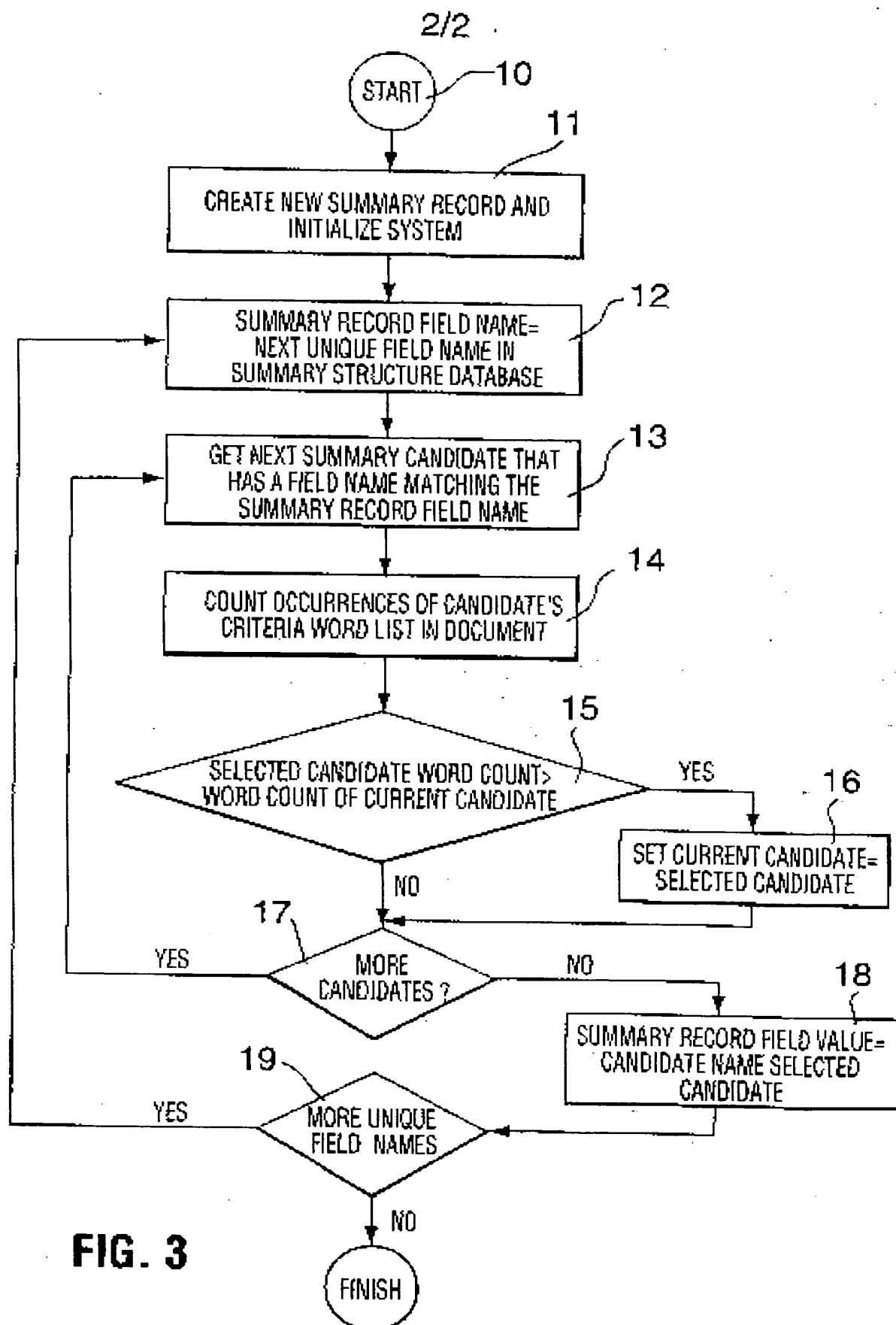


FIG. 2

SUBSTITUTE SHEET (RULE 26)



SUBSTITUTE SHEET (RULE 26)